

5 GTGACCCACGCTCCGGGCGCCCGCGCAGTGTCAGGCTGCAGGCGCGCGGAGGAGGCTGCCGCTCTGGCTTGCCG 79
CCCCCGCGCGCTGCACACGACCGACCCAGCCGCGCTGCCGCGGGCC ATG GAC CTG CCC AGG GGC CTG GTG 8
V A W A L S L W P G F T D T F N M D T R 28
GTG GCC TGG GCG CTC AGC CTG TGG CCA GGG TTC ACG GAC ACC TTC AAC ATG GAC ACC AGG 210
K P R V I P G S R T A F F G Y T V Q Q H 48
AAG CCC CGG GTC ATC CCT GGC TCC AGG ACC GCC TTC TTT GGC TAC ACA GTG CAG CAG CAC 270
5 D I S G N K W L V V G A P L E T N G Y Q 68
GAC ATC AGT GGC AAT AAG TGG CTG TGG GGC GGC CCA CTG GAA ACC AAT GGC TAC CAG 330
K T G D V Y K C P V I H G N C T K L N L 88
AAG ACG GGA GAC GTG TAC AAG TGT CCA GTG ATC CAC GGG AAC TGC ACC AAC CTC AAC CTG 390
G R V T L S N V S E R K D N M R L G L S 108
GGA AGG GTC ACC CTG TCC AAC GTG TCC GAG CGG AAA GAC AAC ATG CGC CTC GGC CTT AGT 450
5 L A T N P K D N S F L A C S P L W S H E 128
CTC GCC ACC AAC CCC AAG GAC AAC AGC TTC CTG GCC TGC AGC CCC CTC TGG TCT CAT GAG 510
C G S S Y Y T T G M C S R V N S N F R F 148
TGT GGG AGC TCC TAC ACC ACA GGG ATG TGT TCA AGA GTC AAC TCC AAC TTC AGG TTC 570
S K T V A P A L Q R C Q T Y M D I V I V 168
TCC AAG ACC GTG GCC CCA GCT CTC CAA AGG TGC CAG ACC TAC ATG GAC ATC GTC ATT GTC 630
L D G S N S I Y P W V E V Q H F L I N I 188
CTG GAT GGC TCC AAC AGC ATC TAC CCC TGG GTG GAG GTT CAG CAC TTC CTC ATC AAC ATC 690
5 L K K F Y I G P G Q I Q V G V Q Y G E 208
CTG AAA AAG TTT TAC ATT GGC CCA GGG CAG ATC CAG GTT GGA GTT GTG CAG TAT GGC GAA 750

FIG. 1A

| | | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| D | V | V | H | E | F | H | L | N | D | Y | R | S | V | K | D | V | V | E | A | 228 |
| GAT | GTG | GTG | CAT | GAG | TTT | CAC | CTC | AAC | GAC | TAC | AGG | TCT | GTA | AAA | GAT | GTG | GTG | GAA | GCT | 810 |
| A | S | H | I | E | Q | R | G | G | T | E | T | R | T | A | F | G | I | E | F | 248 |
| GCC | AGC | CAC | ATT | GAG | CAG | AGA | GGA | GGA | ACA | GAG | ACC | CGG | ACG | GCA | TTT | GGC | ATT | GAA | TTT | 870 |
| A | R | S | E | A | F | Q | K | G | G | R | K | G | A | K | K | V | M | I | V | 268 |
| GCA | CGC | TCA | GAG | GCT | TTT | CAG | AAG | GGT | GGA | AGG | AAA | GGA | GCC | AAG | AAG | GTG | ATG | ATT | GTC | 930 |
| I | T | D | G | E | S | H | D | S | P | D | L | E | K | V | I | Q | Q | S | E | 288 |
| ATC | ACA | GAT | GGG | GAG | TCC | CAC | GAC | AGC | CCA | GAC | CTG | GAG | AAG | GTG | ATC | CAG | CAA | AGC | GAA | 990 |
| R | D | N | V | T | R | Y | A | V | A | V | L | G | Y | Y | N | R | R | G | I | 308 |
| AGA | GAC | AAC | GTA | ACA | AGA | TAT | GCG | GTG | GCC | GTC | CTG | GGC | TAC | TAC | AAC | CGC | AGG | GGG | ATC | 1050 |
| N | P | E | T | F | L | N | E | I | K | Y | I | A | S | D | P | D | D | K | H | 328 |
| AAT | CCA | GAA | ACT | TTT | CTA | AAT | GAA | ATC | AAA | TAC | ATC | GCC | AGT | GAC | CCT | GAT | GAC | AAG | CAC | 1110 |
| F | F | N | V | T | D | E | A | A | L | K | D | I | V | D | A | L | G | D | R | 348 |
| TTC | TTC | AAT | GTC | ACT | GAT | GAG | GCT | GCC | TTG | AAG | GAC | ATT | GTC | GAT | GCC | CTG | GGG | GAC | AGA | 1170 |
| I | F | S | L | E | G | T | N | K | N | E | T | S | F | G | L | E | M | S | Q | 368 |
| ATC | TTC | AGC | CTG | GAA | GGC | ACC | AAC | AAG | AAC | GAG | ACC | TCC | TTT | GGG | CTG | GAG | ATG | TCA | CAG | 1230 |

FIG. 1B

T G F S S H V V E D G G V L L L G A V G A Y 388
 ACG GGC TTT TCC TCG CAC GTG GTG GAG GAT GGG GTT CTG CTG GGA GCC GTC GGT GCC TAT 1290
 D W N G A V L K E T S A G K V I P L R E 408
 GAC TGG AAT GGA GCT GTG CTA AAG GAG ACG AGT GCC GGG AAG GTC ATT CCT CTC CGC GAG 1350
 S Y L K E F P E E L K N H G A Y L G Y T 428
 TCC TAC CTG AAA GAG TTC CCC GAG GAG CTC AAG AAC CAT GGT GCA TAC CTG GGG TAC ACA 1410
 V T S V V S S R Q G R V Y V A G A P R F 448
 GTC ACA TCG GTC GTG TCC TCC AGG CAG GGG CGA GTG TAC GTG GCC GGA GCC CCC CGG TTC 1470
 N H T G K V I L F T M H N N R S L T I H 468
 AAC CAC ACG GGC AAG GTC ATC CTG TTC ACC ATG CAC AAC AAC CGG AGC CTC ACC ATC CAC 1530
 Q A M R G Q Q I G S Y F G S E I T S V D 488
 CAG GCT ATG CGG GGC CAG CAG ATA GGC TCT TAC TTT GGG AGT GAA ATC ACC TCG GTG GAC 1590
 I D G D G V T D V L L V G A P M Y F N E 508
 ATC GAC GGC GAC GGC GTG ACT GAT GTC CTG CTG GGC GCA CCC ATG TAC TTC AAC GAG 1650
 G R E R G G G AAG GTG TAC GTC TAT GAG CTG AGA CAG AAC CGG TTT GTT TAT AAC 528
 GGC CGT GAG CGA GGC AAG GTG TAC GTC TAT GAG CTG AGA CAG AAC CGG TTT GTT TAT AAC 1710
 G T L K D S H S Y Q N A R F G S I A S 548
 GGA ACG CTA AAG GAT TCA CAC CAC AGT TAC CAG AAT GCC CGA TTT GGG TCC TCC ATT GCC TCA 1770
 V R D L N Q CAG GAT TCC TAC AAT GAC GTG GTG GTG GGA GCC CCC CTG GAG GAC 568
 GTT CGA GAC CTC AAC CAG GAT TCC TAC AAT GAC GTG GTG GTG GGA GCC CCC CTG GAG GAC 1830
 N H A G A I Y I F H G F R G S I L K T P 588
 AAC CAC GCA GGA GCC ATC TAC ATC TTC CAC GGC TTC CGA GGC AGC ATC CTG AAG ACA CCT 1890
 K Q R I T A S E L A T G L Q Y F G C S I 608
 AAG CAG AGA ATC ACA GCC TCA GAG CTG GCT ACC GGC CTC CAG TAT TTT GGC TGC AGC ATC 1950

FIG. 1C

| | | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| H | G | Q | L | D | L | N | E | D | G | L | I | D | L | A | V | G | A | L | G | 628 |
| CAC | GGG | CAA | TTG | GAC | CTC | AAT | GAG | GAT | GGG | CTC | ATC | GAC | CTG | GCA | GTG | GGA | GCC | CTT | GGC | 2010 |
| N | A | V | I | L | W | S | R | P | V | V | Q | I | N | A | S | L | H | F | E | 648 |
| AAC | GCT | GTG | ATT | CTG | TGG | TCC | CGC | CCA | GTG | GTT | CAG | ATC | AAT | GCC | AGC | CTC | CAC | TTT | GAG | 2070 |
| P | S | K | I | N | I | F | H | R | D | C | K | R | S | G | R | D | A | T | C | 668 |
| CCA | TCC | AAG | ATC | AAC | ATC | TTC | CAC | AGA | GAC | TGC | AAG | CGC | AGT | GGC | AGG | GAT | GCC | ACC | TGC | 2130 |
| L | A | A | F | L | C | F | T | P | I | F | L | A | P | H | F | Q | T | T | T | 688 |
| CTG | GCC | GCC | TTC | CTC | TGC | TTC | ACG | CCC | ATC | TTC | CTG | GCA | CCC | CAT | TTC | CAA | ACA | ACA | ACT | 2190 |
| V | G | I | R | Y | N | A | T | M | D | E | R | R | Y | T | P | R | A | H | L | 708 |
| GTT | GGC | ATC | AGA | TAC | AAC | GCC | ACC | ATG | GAT | GAG | AGG | CGG | TAT | ACA | CCG | AGG | GCC | CAC | CTG | 2250 |
| D | E | G | G | D | R | F | T | N | R | A | V | L | L | S | S | G | Q | E | L | 728 |
| GAC | GAG | GGC | GGG | GAC | CGA | TTC | ACC | AAC | AGA | GCC | GTA | CTG | CTC | TCC | TCC | GGC | CAG | GAG | CTC | 2310 |
| C | E | R | I | N | F | H | V | L | D | T | A | D | Y | V | K | P | V | T | F | 748 |
| TGT | GAG | CGG | ATC | AAC | TTC | CAT | GTC | CTG | GAC | ACT | GCT | GAC | TAC | GTG | AAG | CCA | GTG | ACC | TTC | 2370 |
| S | V | E | Y | S | L | E | D | P | D | H | G | P | M | L | D | D | G | W | P | 768 |
| TCA | GTC | GAG | TAT | TCC | CTG | GAG | GAC | CCT | GAC | CAT | GGC | CCC | ATG | CTG | GAC | GAC | TGG | CCC | | 2430 |

FIG. 1D

T T L R V S V P F W N G C N E D E H C V 788
 ACC ACT CTC AGA GTC TCG GTG CCC TTC TGG AAC GGC TGC AAT GAG GAT GAG CAC TGT GTC 2490

 P D L V L D A R S D L P T A M E Y C Q R 808
 CCT GAC CTT GTG TTG GAT GCC CGG AGT GAC CTG CCC ACG GCC ATG GAG TAC TGC CAG AGG 2550

 V L R K P A Q D C S A Y T L S F D T T V 828
 GTG CTG AGG AAG CCT GCG CAG GAC TGC TCC GCA TAC ACG CTG TCC TTC GAC ACC ACA GTC 2610

 F I I E S T R Q R V A V E A T L E N R G 848
 TTC ATC ATA GAG AGC ACA CGC CAG CGA GTG GCG GTG GAG GCC ACA CTG GAG AAC AGG GGC 2670

 E N A Y S T V L N I S Q S A N L Q F A S 868
 GAG AAC GCC TAC AGC ACG GTC CTA AAT ATC TCG CAG TCA GCA AAC CTG CAG TTT GCC AGC 2730

 L I Q K E D S D G S I E C V N E E R R L 888
 TTG ATC CAG AAG GAG GAC TCA GAC GGT AGC ATT GAG TGT GTG AAC GAG GAG AGG AGG CTC 2790

 Q K Q V C N V S Y P F F R A K A K V A F 908
 CAG AAG CAA GTC TGC AAC AAC GTC AGC TAT CCC TTC TCG CGG GCC AAG GGC AAC GTG GCT TTC 2850

 R L D F E F S K S I F L H L E I E L A 928
 CGT CTT GAT TTT GAG TTC AGC AAA TCC ATC TTC CTA CAC CAC CTG GAG ATC GAG CTC GCT 2910

 A G S D S N E R D S T K E D N V A P L R 948
 GCA GGC AGT GAC AGT AAT GAG CGG GAC AGC ACC AAG GAA GAC AAC GTG GCC CCC TTA CGC 2970

 F H L K Y E A D V L F T R S S L S H Y 968
 TTC CAC CTC AAA TAC GAG GCT GAC GTC CTC TTC ACC AGG AGC AGC CTG AGC CAC TAC 3030

 E V K P N S S L E R Y D G I G P P F S C 988
 GAG GTC AAG CCC AAC AGC TCG CTG GAG AGA TAC GAT GGT ATC GGG CCT CCC TTC AGC TGC 3090

 I F R I Q N L G L F P I H G M M K I T 1008
 ATC TTC AGG ATC CAG AAC TTG GGC TTG TTC CCC ATC CAC GGG ATG ATG ATG AAG ATC ACC 3150

FIG. 1E

I P I A T R S G N R L L L K L R D F L T D 1028
 ATT CCC ATC GCC ACC AGG AGC GGC AAC CGC CTA CTG AAG CTG AGG GAC TTC CTC ACG GAC 3210

 E A N T S C N I W G N S T E Y R P T P V 1048
 GAG GCG AAC ACG TCC TGT AAC ATC TGG GGC AAT AGC ACT GAG TAC CGG CCC ACC CCA GTG 3270

 E E D L R R A P Q L N H S N S D V V S I 1068
 GAG GAA GAC TTG CGT CGT GCT CCA CAG CTG AAT CAC AGC AAC TCT GAT GTC GTC TCC ATC 3330

 N C N I R L V P N Q E I N F H L L G N L 1088
 AAC TGC AAT ATA CGG CTG GTC CTC CCC AAC CAG GAA ATC AAT TTC CAT CTA CTG GGG AAC CTG 3390

 W L R S L K A L K Y K S M K I M V N A A 1108
 TGG TTG AGG TCC CTA AAA GCA CTC AAG TAC AAA TCC ATG AAA ATC ATG GTC AAC GCA GCC 3450

 L Q R Q F H S P F I F R E E D P S R Q I 1128
 TTG CAG AGG CAG TTC CAC AGC CCC TTC ATC TTC CGT GAG GAG GAT CCC AGC CGC CAG ATC 3510

 V F E I S K Q E D W Q V P I W I I V G S 1148
 GTG TTT GAG ATC TCC AAG CAA GAG GAC TGG CAG GTC CCC ATC TGG ATC ATT GTA GGC AGC 3570

 T L G G L L L L A L L V L A L W K L G F 1168
 ACC CTG GGC GGC CTC CTA CTG CTG GCC CTG CTG GTC GTC GCA CTG TGG AAG CTC GGC TTC 3630

 F R S A R R R R E P G L D P T P K V L E 1188
 TTT AGA AGT GCC AGG CGC AGG AGG GAG CCT GGT CTG GAC CCC ACC CCC AAA GTG CTG GAG 3690

S * 1189
 TGA 3693

FIG. 1F

GGCTCCAGAGGAGACTTTGAGTTGATGGGGGCCAGGACACACCAGTCCAGGTAGTGTGAGACCCAGGCCCTGTGGCCCCCACC 3772
CGAGCTGGAGCGGAGAGCCAGCTGGCTTTGCACTTGACCTCATCTCCCGAGCAATGGCGCCTGCTCCCTCCAGAA 3851
TGGAACCTCAAGCTGGTTTAAAGTGGAACCTGCCCTACTGGGAGACTGGGACACCTTTAACACAGACCCCTAGGGATTAA 3930
AGGGACACCCCTACACACACCCAGGCCACGCCCAAGGCCTCCCTCAGGCTCTGTGGAGGGCATTTGCTGCCCCCAGCTAC 4009
TAAGGTGCTAGGAATTCGTAATCATCCCCATCCTCCAGAGAAACCCAGGGAGGAAGACTGTAAATACGAACCCCAATCTG 4088
CACACTCCAGGCCCTCTAGTTCCAGAAGGATCCAAGACAAAACAGATCTGAAATTCTGCCCTTTTCTCTCACCCCATCCCCAC 4167
CCCTCCATTGGCTCCCCAAGTCACACCCCACTCCCTTCCCCATAGATAGGCCCTCTGGGGCTCCTGAAGAAATGAACCCCAAGA 4246
GCAAGGGCTTGATGGTGACAGCTGCAAGCCAGGGATGAAGAAAGACTCTGAGATGTGGAGACTGTATGGCCAGGCAAGTG 4325
GGACCAGGATACTGGACGCTGTCCTGAGATGAGAGGTAGCCGGGCTCTGCACCCACGTGCATTCACATTGACCCGCAACT 4404
CACACATTCCCCCACCAGCTGCAGCCCCCTTGCTCTCAGCTGCCAACCCCTCCCGGGTCACTTTTGTTCGCCAGGTACCTCA 4483
TGGGAAGCATGTGGATGACACAAATCCCTGGGGCTGTGCATTCCCCACGTCTTCTTGCTGCAGCCCTGCCCCCTAGACATGGA 4562
CGCACTGGCCTGGCTGCAGCTGGGCAGCAGGGGTAGGGGTAGGGAGCCTCCCCCTCCCTGTATCACCCCTCCCTACACA 4641
CACACACACACACACACACTGCCCTCCCATCCTTCCCTCATGCCCCGCCAGTGCACAGGGAAGGGCTTGGCCAGC 4720
GCTGTTGAGGGTCCCCCTCTGGAATGCACTGAATAAAGCACCGTGCAAGGACTCCCGGAGCCTGTGCAGCCTTGGTGGCA 4799
AATATCTCATCTGCCCGGCCCCCAGGACAAAGTGTATGACCAGTGATAATGCCCCCAAGGACAAAGGGCGTGCCTGGCGCC 4878
CAGTGGAGTAATTATGCCTTAGTCTTGTTTGGAGGTAGAAATGCAAGGGGGACACATGAAAGGCATCAGTCCCCCTGT 4957
GCATAGTACGACCTTTACTGTCGTATTTTGAATAATTAATAATACAGTGTTTAAAAACAATAAAAAAAGGGC 5036
GGCCGC 5042

FIG. 1G

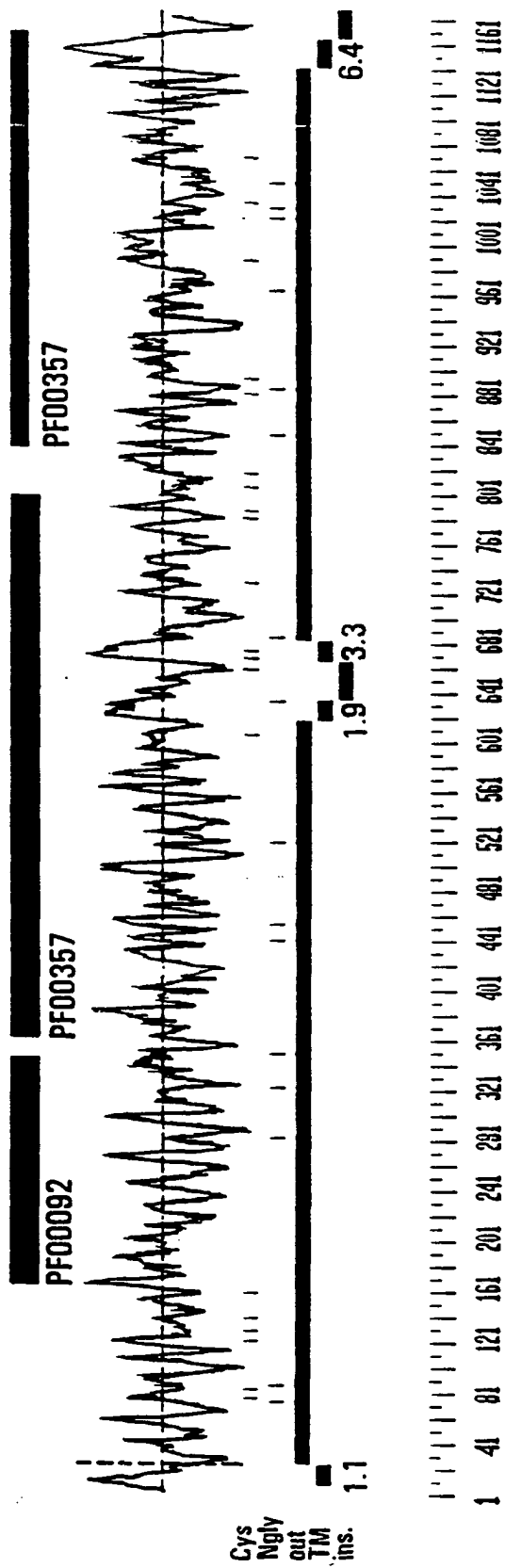


FIG. 2


```

      590      700      710      720      730      740      750
inputs ATCAACATCCTGAAAAAGTTTTACATTGGCCCAGGGCAGATCCAGGTTGGAGTTGTGCAGTATGGCGAAG
      590      600      610      620      630      640      650
      GAAGACTGGTAGGGAACTGTTT-ATTGACCCAGAACAGATACAGGTGGGACTGGTACAGTATGGGGAGA

      760      770      780      790      800      810      820
inputs ATGTGGTGCATGAGTTTTACCTCAACGACTACAGGTCTGTAAAAGATGTGGTGAAGCTGCCAGCCACAT
      660      670      680      690      700      710      720
      GCCCTGTACATGAGTGGTCCCTGGGAGATTTCCGAACGAAGGAAGAAGTGGTGAGAGCAGCAAAGAACCT

      830      840      850      860      870      880
inputs TGAGCAGAGAGGAGGAACAGAGACCCGACGGCATTGGCATTGAATTTGCACGCTCAGAGGCTTTC---
      730      740      750      760      770      780      790
      CAGTCGGCGGGAGGGACGAGAAACAAAGACTGCCCAAGCAATAATGGTGGCCTGCACAGAAGGGTTTCAGT

      890      900      910      920      930      940      950
inputs CAGA---AGGGTGAAGGAAAGGAGCCAAGAAGGTGATGATTGTCATCACAGATGGGGAGTCCCACGACA
      800      810      820      830      840      850      860
      CAGTCCCATGGGGCCGACCCGAGGCTGCCAGGCTACTGGTGGTTGTCACTGATGGAGAGTCCCATGATG

      960      970      980      990      1000      1010      1020
inputs GCCCAGA---CCTGGAGAA-GGTGATCCAGCAAAGCGAAAGAGACAACGTAACAAGATATGCGGTGGCC
      870      880      890      900      910      920      930
      GAGAGGAGCTTCCTGCAGCACTAAAGGCCTGTGAGGC-TGGAAGA---GTGACACGCTATGGGATTGCA

      1030      1040      1050      1060      1070      1080      1090
inputs GTCCTGGGCTACTACAACCGCAGGGGGATCAATCCAGAACTTTCTAAATGAAATCAAATACATCGCCA
      940      950      960      970      980      990      1000
      GTCCTTGGTCACTACCTCCGGCGGCAGCGAGATCCAGCTCTTTCCTGAGAGAAATTAGAACTATTGCCA

      1100      1110      1120      1130      1140      1150      1160
inputs GTGACCCTGATGACAAGCACTTCTTCAATGTCACCTGATGAGGCTGCCTTGAAGGACATTGTGATGCCCT
      1010      1020      1030      1040      1050      1060      1070
      GTGATCCAGATGAGCGATTCTTCTTCAATGTCACAGATGAGGCTGCTCTGACTGACATTGTGGATGCACT

      1170      1180      1190      1200      1210      1220
inputs GGGGGACAGAATCTTTCAGCCTGGAAGGCACC--AACA-AGAACGAGACCTCCTTTGGGCTGGAGATGTCA
      1080      1090      1100      1110      1120      1130      1140
      AGGAGATCGGATTTTGGCCTTGAAGGTCCCATGCAGAAAACGAAAGCTCCTTTGGGCTGGAAATGTCT

      1230      1240      1250      1260      1270      1280      1290
inputs CAGACGGGCTTTTTCCTCGCACGTGGTGGAGGATGGGGTTCTGCTGGGAGCCGTCGGTGCCTATGACTGGA
      1150      1160      1170      1180      1190      1200      1210
      CAGATTGGTTTCTCCACTCATCGGCTAAAGGATGGGATTCTTTTGGGATGGTGGGGCCCTATGACTGGG

      1300      1310      1320      1330      1340      1350      1360
inputs ATGGAGCTGTGCTAAAG-GAGACGAGTGCCGGGAAGGTCATTCTCTCCGCGAGTCCTACCTGAAA--GA
      1220      1230      1240      1250      1260      1270
      GAGGCTCTGTGCTATGGCTTGAAGGAGGCC---ACCGCCTTTTCCCCCACGAATGGCACTGGAAGACGA

```

FIG. 3B


```

      2120      2130      2140      2150      2160      2170      2180
inputs  AGGGATGCCACCTGCCTGGCCGCTTCTCTGCTTCACGCCCATCTTCTGGCACCCCATTTTC-CAAACA
      ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      CAAGAAGCAGTCTGTCTGACTGCAGCCCTTTGCTTC-CAAGTGACCTCCCGT-CTCCTGGTCTGCTGGGA
      2040      2050      2060      2070      2080      2090

      2190      2200      2210      2220      2230      2240      2250
inputs  ACAACTGTT-GGCATCAGATACAACGCCACCATGGATGAGAGG-CGGTATACAC-CGAG-GGCCCACCTG
      ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      TCACCAATTCTACATGAGGTTTCACCGCATCACTGGATGAATGGACTGCTGGGGCACGTGCAGCATTGAT
      2100      2110      2120      2130      2140      2150      2160

      2260      2270      2280      2290      2300      2310
inputs  GACGAGGGCGGGGACCGATTACCAACAGAGCCGTACTGCTCTCCTCCGGCCAGGAGCTCT-GTGAGCGG
      ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      GGCTCTGGCCAGAG--GTTGTCCCTCGGAGGC-TCCGGCTCAG-TGTGGGAATGTCATTGTGAGCAG
      2170      2180      2190      2200      2210      2220      2230

      2320      2330      2340      2350      2360      2370      2380
inputs  ATCAACTTCCATGTCTTGACACTGCTGACTACGTGAAGCCAGTGACCTTCTCAGTCGAGTATTCCTGG
      ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      CTACACTTCCATGTGCTGGATACATCAGATTACCTCCGGCCAGTGGCCTTGACTGTGACCTTTGCCTTGG
      2240      2250      2260      2270      2280      2290      2300

      2390      2400      2410      2420      2430      2440      2450
inputs  A--GGAC--CCTGACCATGGCCCCATGCTGGACGACGGCTGGCCCACTCT-CAGAGTCTCGGTGCCC
      ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      ACAATACTACAAAGCCAGGGCCT-GTGCTGAATGAGGGCTCACCCACCTCTATACAAAAGCT-GGTCCCC
      2310      2320      2330      2340      2350      2360      2370

      2460      2470      2480      2490      2500      2510      2520
inputs  TTCTGGAACGGCTGCAATGAGGATGAGCACTGTGTCCCTGACCTTGTGTTGGATGCCCGGAGTGACCTGC
      ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      TTCT-----CAA--AGGAT-----TGTGGCCCTGACAATGAAT--GTGTAC-----AGACCTG-
      2380      2390      2400      2410

      2530      2540      2550      2560      2570      2580      2590
inputs  CCACGGCCATGGAGTACTGCCAGAGGGTGTGAGGAAGCCTGCGCAGGACTGCTCCGCATACACGCTGTC
      ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      -----GTGCTTCAAGTGAA--TATGGA--CATCAGAGG--CTCC-----AGGAAGGC
      2420      2430      2440      2450

      2600      2610      2620      2630      2640      2650      2660
inputs  CTTCGACACCACAGTCTTCATCATAGAGAGCACACGCCAGCGAGTGGCGGTGGAGGCCACACTGGAGAAC
      ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      C-----CCA-TTTGTGGTTCGAGGTG-GC---CGGCGGAAAGTGCTGGTATCTACAACTCTGGAGAAC
      2460      2470      2480      2490      2500      2510

      2670      2680      2690      2700      2710      2720      2730
inputs  AGGGGCGAGAACGCCTACAGCACGGTCTAAATATCTCGCAGTCAGCAAACCTGCAGTTTGCCAG-CTTG
      ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      AGAAAGGAAAATGCTTACAATACGAGCCTGAGTATCATCTTCTCTAGAAACCTCCACCTGGCCAGTCTCA
      2520      2530      2540      2550      2560      2570      2580

      2740      2750      2760      2770      2780      2790      2800
inputs  ATCC--AGAAGGAGGACTCAGACGGTAGCATTGAGTGTGT-GAACGAGGAGAGGAGGCTCCAGAAGCAAG
      ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      CTCTCAGAGAGAGAGCCCAAT--AAAG-GTGGAATGTGCCGCCCTTCTGCTCATGC-CCG---GC---
      2590      2600      2610      2620      2630      2640

      2810      2820      2830      2840      2850      2860      2870
inputs  TCTGCAACGTCAGCTATCCCTTCTTCCGGGCCAAGGCCAAGGTGGCTTTCCGTCTTGATTTGAGTTTACG
      ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      TCTGCACTGTGGGGCATCCTGTCTTCCAGACTGGAGCCAAGGTGACCTTCTGCTAGAGTTTGTAGTTTACG
      2650      2660      2670      2680      2690      2700      2710

```

FIG. 3D

```

      2880      2890      2900      2910      2920      2930
inputs CAAATCCATCT-TCCTACACCACCTGGAGATCGAGCTCGCTGCAGGCAGTGACAGTAATGAGCGGGACAG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      CTGCTCC-TCTCTCTGAGCCAGGTCCTTTGGGAAGCTGACTGCCAGCAGTGACAG-----
      2720      2730      2740      2750      2760      2770

      2940      2950      2960      2970      2980      2990      3000
inputs CACCAAGGAAGACAACGTGGCCCCCTTACGCTTCCACCTCAAATACGAGGCTGACGTCTCTTACCAGG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      --CCTGGAGAGAAA--TGGCACCCTT-----CAAGAAAAC-----ACAGCC-----CAGA
      2780      2790      2800      2810

      3010      3020      3030      3040      3050      3060      3070
inputs AGCAGCAGCCTGAGCCACTACGAGGTCAAGCCCAACAGCTCGCTGGAGAGATACGATGGTATCGGGCCTC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      --CCTCAGCCT-----ACATCCAATATGAGCCCCAC--CTC-CTG-----TTCTCTAGT---GAGTCTA
      2820      2830      2840      2850      2860

      3080      3090      3100      3110      3120      3130      3140
inputs CCTTCAGCTGCATCTTCAGGATCCAGAACTTGGGCTTGTTCCCCATCCACGGGATGATGATGAAGATCAC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      CC-----CTGCA-----CC-----GCT-----ATGAGTTTCAC
      2870      2880

      3150      3160      3170      3180      3190      3200      3210
inputs CATTTCCATCGCCACCAGGAGCGGCAACCGCCTACTGAAGCTGAGGGACTTCCTCACGGACGAGGCGAAC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      C-----CATATGGGAC-----CCTCC-----CAGTGGG--TCCT-----GGC-----
      2890      2900      2910

      3220      3230      3240      3250      3260      3270      3280
inputs ACGTCCTGTAACATCTGGGGCAATAGCACTGAGTACCGGGCCACCCAGTGGAGGAAGACTTGGCTCGTG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      ---CCAGAAT--TCAAAACCACT--CTCAGGGTTCAGAACCT-----AGGCT-GCTATGTG
      2920      2930      2940      2950      2960

      3290      3300      3310      3320      3330      3340      3350
inputs CTCCACAGCTGAATCACAGCAACTCTGATGTCTGCTCCATCAACTGCAATATACGGCTGGTCCCCAACCA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      GTCAGTGGC-----CTCATCATCTCA---GCC--CTCCTTC--CAGCTGT---GGC-----CCA
      2970      2980      2990      3000

      3360      3370      3380      3390      3400      3410      3420
inputs GGAAATCAATTTCCATCTACTGGGGAACTGTGGTTGAGGTCCCTAAAAGCACTCAAGTACAAATCCATG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      TGGGGGCAATTAC-----TTCCTAT-----CACTGT-----CTCAAGT-CATCAC-----
      3010      3020      3030      3040

      3430      3440      3450      3460      3470      3480      3490
inputs AAAATCATGGTCAACGCAGCCTTGCGAGAGGCAGTTCACAGCCCCCTTCATCTTCCGTGAGGAGGATCCCA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      --TAACAATG-CAA-GCTGCATAG---TGCAGAACCT-----GACT-----GAACCC-
      3050      3060      3070      3080

      3500      3510      3520      3530      3540      3550      3560
inputs GCCGCCAGATCGTGTTTGAGATCTCCAAGCAAGAGGACTGGCAGGTCCCCATCTGGATCATTTGTAGGCAG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      -----CCAGGC-----CCACCTG-----TG
      3090      3100

      3570      3580      3590      3600      3610      3620      3630
inputs CACCCTGGGGGGCCTCCTACTGCTGGCCCTGCTGGTCTGGCACTGTGGAAGCTCGGCTTCTTTAGAAGT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      CATCCAGAGGAGCTTC-----AACACA-CAAA--CAGACTGAATGGGAGCAATACT-C-----AGT
      3110      3120      3130      3140      3150

```

FIG. 3E


```

4410      4420      4430      4440      4450      4460      4470
inputs ATTCCCCCACCAGCTGCAGCCCCCTTGCTCTCAGCTGCCAACCCTCCCGGGTCACTTTTGTTCACAGGTA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      TATCTCCC-----CAGACC-----AGCAGCCTGACTT-----G-ACCTTTGA-----GT-
              3640              3650              3660              3670

4480      4490      4500      4510      4520      4530      4540
inputs CTCATGGGAAGCATGTGGATGACACAATCCCTGGGGCTGTG-CATTCCCACGTCTTCTTGCTGCAGCCT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      CT--AGGGATGCTGCTGG-----CTAGAGATGAGGCTTTACCTCAG--ACAAGAAG-AGCTG
              3680              3690              3700              3710              3720

4550      4560      4570      4580      4590      4600      4610
inputs CCCCTAGACATGGACGCACTGGCCTGGCTGCAGCTGGGCAGCAGGGGTAGGGGTAGGGAGCCTCCCCCTC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      CACCAAAAC-----TAGCCATGCT-----CCCAC
              3730              3740

4620      4630      4640      4650      4660      4670      4680
inputs CTGTATCACCCCTCCCTACACACACACACACACACACACACACTGCCTCCCATCCTTCCTCAT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      CT-----CTGCTTCCC-TCCT--CCTCGT
              3750              3760              3770

4690      4700      4710      4720      4730      4740      4750
inputs CCCGCCAGTGCACAGGGAAGGGCTTGGCCAGCGCTGTTGAGGGGTCCCCCTCTGGAATGCACTGAATAAA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      ATC-CTGGTTC-----CATAGCCAACACTG-----GGG-----CTTTTGTMTG-----
              3780              3790              3800

4760      4770      4780      4790      4800      4810      4820
inputs CACGTGCAAGGACTCCCGGAGCCTGTGCAGCCTTGGTGGCAAATATCTCATCTGCCGGCCCCCAGGACA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      -----GGGTCCTTT-----TATC-----CCCAGGA--
              3810              3820

4830      4840      4850      4860      4870      4880      4890
inputs GTGGTATGACCAGTGATAATGCCCCAAGGACAAGGGGCGTGCCTGGCGCCAGTGGAGTAATTTATGCC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      -----ATCAATAATT-----TTTTTGCC
              3830              3840

4900      4910      4920      4930      4940      4950      4960
inputs TAGTCTTGTGTTTGGAGGTAGAAATGCAAGGGGGACACATGAAAGGCATCAGTCCCCCTGTGCATAGTACC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      -----AGGAAAAAA-----AAAAGC-----GGCC-----GCG-AATTCC
              3850              3860              3870

4970      4980      4990      5000      5010      5020      5030
inputs CCTTTACTGTCGTATTTTTGAAAAATTAAAAATACAGTGTTTAAAAACAAAAAAGGGGCC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      -----TATC-----AAGCT-----
              3880

5040
inputs CCGC
      ----

```

FIG. 3G

```

      10      20      30      40      50      60      7
inputs MDLPRGLVVAWALSLWPGFTDTFNM DTRKPRVIPGSRTAFFGYTVQQHDISGNKWL VVGAPLETNGYQK
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      MELPFVTHLFLPLVFLTGLCSPFNLD EHHPRLEFGPPEAEFGYSVLQHVGGGQRWMLV GAFWDGPGSGDR
      10      20      30      40      50      60      7

      80      90      100      110      120      130
inputs GDVYKCPV--IHGN-CTKLN LGRVTL SNVSEKDNMRLGLSLATNPKD NSFLAC SPLWSHECGSSYYT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      GDVYRCFVGGAHNAPCAK GHLGDYQLGNSSHPAVNMHLGMSLLET DGDGGF MACAPLWSRACGSSVFS
      80      90      100      110      120      130      14

      140      150      160      170      180      190      200
inputs MCSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDG SNSIYPWVEVQHFLINILKKFYIGPGQIQVGVVQ
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      ICARVDASFQPGSLAPTAQRCPTYMDVVIVLDG SNSIYPWSEVQTF LRLVLGKLFIDPEQIQVGLVQ
      150      160      170      180      190      200      21

      210      220      230      240      250      260      270
inputs EDVVHEFHLNDYRSVKDVVEAASHIEQRGGTETRTAFGIEFARSEAFQK--GGRKGAKKVMIVITDGE
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      ESPVHEWSLGD FRTKEEVVRAAKNLSRREGRET KTAQAIMVACTEGFSQSHGGRPEAARLLVVVTDGE
      220      230      240      250      260      270      28

      280      290      300      310      320      330      340
inputs DSPDLEKVIQQSERDNVTRYAVAVLGYYNRRGINPETFLNEIKYIASDPDDKHFFNVTD EAALKDIVD
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      DGEELPAALKACEAGRVTRYGIAVLGHYLRQRDPSSFLREIRTIASDPDERFFFNVTD EAALTDIVD
      290      300      310      320      330      340      35

      350      360      370      380      390      400      410
inputs GDRIFSLEGTN-KNETSFGLEMSQTGFSSHVVEDGVLLGAVGAYDWN GAVLKETSAGKVIPLRESYLK
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      GDRIFGLEGSHAENESSFGLEMSQIGFSTHRLKDGILFGMVGAYDWGGSVLWLEGGHRLFP PMALED
      360      370      380      390      400      410      42

      420      430      440      450      460      470      480
inputs PEELKNHGAYLG YTVTSVSSRQGRVYVAGAPRFNHTGKVILFTMHN NRSLTIHQAMRGQQIGSYFGS
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      PPALQNHAAYLGYSVSSMLLRGGRRLLFSGAPRFHRGKVIAFQLKKDGAVRVAQSLQGEQIGSYFGS
      430      440      450      460      470      480      49

      490      500      510      520      530      540      550
inputs TSVDIDGDGVTDVLLVGAPMYFN-EGRERGKVYVYEL-RQNR FVYNGTLKDSHSYQNARFGSSIASVR
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      CPLDTRDGT TDVLLVAAPMFLGFPQNKETGRVYVYLVGQQSLLTLQGT LQPEPP-QDARFGFAMGALP
      500      510      520      530      540      550

      560      570      580      590      600      610      620
inputs NQDSYNDVVVGAPLEDNHAGAIYIFHGFRGSILKTPKQRITASELATGLQYFGCSIHGQLDLNEDGLI
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      NQDGFADVAVGAPLEDGHQ GALYLYHGTQSGVRPHPAQRIAAASMPHALSYFGRSVDGRLLDGDGDLV
      560      570      580      590      600      610      620

      630      640      650      660      670      680      690
inputs AVGALGNAVILWSRPVVQINASLHFEP SKINIFHRDCKRSGRDATCLAAFLCFTPIFLAPHFQTTTVG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      AVGAQGAAILLSSRPVHLTPSLEVTPQAISVVQRDCRRRGQEAVCLTAALCFQVTSRTPGRWDHQFY
      630      640      650      660      670      680      690

```

FIG. 4A


```

      700      710      720      730      740      750      760
inputs YNATMDERRYTPRAHLDEGGDRFTNRAVLLSSGQELCERINFHVLDTADYVKPVTFSVEYSLEDPD-I
      .....
      FTASLDEWTAGARAAFDGSGQRLSPRRLRLSVGNVTCEQLHFHVLDTS DYLRPVALTVTFALDNTTKI
      700      710      720      730      740      750      760

      770      780      790      800      810      820      830
inputs MLDDGWPTTLRVSVPFWNGCNEDEHCVDPDLVLDARSDLP TAMEYCQRVLRKPAQDCSAYTLSFDTTV
      .....
      VLNEGSPTSIOQLVPFSKDCGPDNECVTDLV LQVNM DI-----RGSRKAP-----
      770      780      790      800      810

      840      850      860      870      880      890      900
inputs ESTRQRVAVEATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNEERRLQKQVCNVSYPI
      .....
      RGGRRKVLVSTTLENRKENAYNTSLSIIFSRNLHLASLT PQRESPIKVECAAPSA--HARLCSVGHP
      820      830      840      850      860      870      880

      910      920      930      940      950      960      970
inputs AKAKVAFRLDFEFSKSIFLHHLEIELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLFTRSSSLSHYI
      .....
      TGAKVTFLLEFEFSCSSLLSQVFGKLTASSDSLERNGT LQENTAQT SAYIQYEPHLLFSSESTLHRYI
      890      900      910      920      930      940      950

      980      990      1000      1010      1020      1030      1040
inputs PNSSLERYDGIGPPFSCIFRIQNLGLFPIHGMMKITIPIATRSNGNRLKL RDLTDEANTSCNIWGI
      .....
      PYGTLPV--GPGPEFKTTLRVQNLGCVVSGLIISALLPAVAHGGNYFLSLSQVITN--NASCIV-QI
      960      970      980      990      1000      1010      1020

      1050      1060      1070      1080      1090      1100
inputs EYRPTPVE-EDLRRAPQLNHSNSDVVSINCNI-RLVPNQEINFHLLGNLWLRSLKALKYKSMKIMVN
      .....
      EPPGPPVHPEELQHTNRLNGSNTQCQVVRCHLGQLAKGTEVSVGLLRLVHNEFFRRAKFKSLTVVSTI
      1030      1040      1050      1060      1070      1080      1090

      1110      1120      1130      1140      1150      1160      1170
inputs QRQFHSPFIFREEDPSRQIVFEISKQEDWQVPIWIIVGSTLGGLLLLALLVLALWKLGFFRSARRRR
      .....
      GTEEGSVLQLTEASRWSESLLLEVQTRPILISLWILIGSVLGGLLLLALLVFC LWKLGFF--AHKKII
      1100      1110      1120      1130      1140      1150

      1180
inputs LDPTPKVLE
      .....
      EKREEKLEQ
      1160

```

FIG. 4B

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|-----|
| GGGACACCCGCGCGCGCGCGG | M | D | F | P | R | G | L | L | V | A | W | T | L | 13 |
| ATG GAC TTT CCC AGG GGC CTC CTG GTG GCC TGG ACC CTC | | | | | | | | | | | | | | 66 |
| S L W P G F T D T F N M D T R N P R V I | | | | | | | | | | | | | | 33 |
| AGC CTT TGG CCA GGA TTC ACG GAC ACC TTC AAT ATG GAT ACC AGA AAT CCC CGG GTC ATT | | | | | | | | | | | | | | 126 |
| A G P S A A F F G Y T V Q Q H D I S G K | | | | | | | | | | | | | | 53 |
| GCT GGT CCC AGT GCT GCT GCC TTC TTT GGC TAC ACA GTA CAG CAT GAT ATC AGT GGC AAG | | | | | | | | | | | | | | 186 |
| K W L V V G A P M E T N G H Q K T G D V | | | | | | | | | | | | | | 73 |
| AAG TGG TTG GTT GTG GGT GCC CCA ATG GAA ACC AAT GGC CAT CAG AAG ACA GGA GAC GTA | | | | | | | | | | | | | | 246 |
| Y K C P V T Q G N C T K L N L G R V T L | | | | | | | | | | | | | | 93 |
| TAC AAA TGT CCG GTA ACC CAG GGC AAC TGC ACC AAG CTC AAG CTG GGC AGG GTC ACT CTG | | | | | | | | | | | | | | 306 |
| S N V S E R K D N M R L G L S L A T N P | | | | | | | | | | | | | | 113 |
| TCC AAT GTG TCT GAG AGG AAG GAC AAC ATG CGC CTC GGC CTG AGC CTT GCC ACC AAC CCC | | | | | | | | | | | | | | 366 |
| K D N S F L A C S P L W S H E C G S Y | | | | | | | | | | | | | | 133 |
| AAG GAC AAC AGC TTC CTG GCC TGC AGC CCT CTC TGG TCG CAC GAG TGT GGA AGC TCC TAC | | | | | | | | | | | | | | 426 |
| Y T T G M C S R V N S N F R F S K T V A | | | | | | | | | | | | | | 153 |
| TAC ACC ACT GGC ATG TGC TCA CGG GTC AAC TCC AAC TTC AGA TTC TCT AAG ACG GTA GCT | | | | | | | | | | | | | | 486 |
| P A L Q R C Q T Y M D I V I V L D G S N | | | | | | | | | | | | | | 173 |
| CCG GCA CTT CAG AGG TGC CAG ACT TAC ATG GAC ATC GTC ATT GTC TTA GAT GGC TCC AAC | | | | | | | | | | | | | | 546 |
| S I Y P W V E V Q H F L I N I L K K F Y | | | | | | | | | | | | | | 193 |
| AGC ATC TAC CCC TGG GTG GAG GTC CAA CAC TTC CTC ATC AAT ATC CTC AAA AAG TTC TAC | | | | | | | | | | | | | | 606 |
| I G P G Q I Q V G I V Q Y G E D A V H E | | | | | | | | | | | | | | 213 |
| ATT GGC CCC GGC CAG ATC CAG GTC GGA ATA GTC CAG TAT GGA GAA GAT GCC GTC CAT GAG | | | | | | | | | | | | | | 666 |

FIG. 5A

| | | | | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|-----|
| F | H | L | N | D | GAC | TAC | Y | R | S | V | K | D | V | V | E | A | A | S | H | I | E | 233 |
| TTC | CAC | CTT | AAT | GAC | TAC | AGG | TCT | AGG | TCT | GTA | AAA | GAT | GTG | GTG | GAA | GCC | GCC | AGC | CAC | ATT | GAG | 726 |
| Q | R | G | G | T | E | T | R | T | A | T | A | F | G | I | E | F | A | R | S | E | A | 253 |
| CAG | AGA | GGA | GGG | ACA | GAG | ACC | CGC | ACG | GCA | TTT | GGC | ATT | GAA | TTT | GCA | CGC | TCG | GAG | GCT | | 786 | |
| F | Q | K | G | G | R | K | G | A | K | K | V | M | I | V | I | T | D | G | E | | 273 | |
| TTC | CAG | AAG | GGT | GGA | AGA | AAA | GGG | GCC | AAG | AAA | GTG | ATG | ATT | GTC | ATC | ACG | GAC | GGG | GAA | | 846 | |
| S | H | D | S | P | D | L | E | K | V | I | R | Q | S | E | K | D | N | V | T | | 293 | |
| TCC | CAC | GAC | AGC | CCA | GAC | CTG | GAG | AAG | AAG | GTG | ATC | CGG | CAG | AGC | GAG | AAG | GAC | AAC | GTG | ACC | 906 | |
| R | Y | A | V | A | V | L | G | Y | Y | Y | N | R | R | G | I | N | P | E | T | F | 313 | |
| AGA | TAC | GCT | GTG | GCC | GTT | TTG | GGC | TAC | TAC | TAC | AAC | CGC | AGG | GGG | ATC | AAT | CCA | GAG | ACT | TTT | 966 | |
| L | N | E | I | K | Y | I | A | S | D | P | D | D | D | K | H | F | N | V | T | | 333 | |
| CTA | AAT | GAA | ATC | AAA | TAC | ATC | GCC | AGC | AGC | GAC | CCT | GAC | GAC | AAG | CAC | TTC | TTC | AAC | GTC | ACA | 1026 | |
| D | E | A | A | L | K | D | I | V | D | A | L | G | D | R | I | F | S | L | E | | 353 | |
| GAT | GAG | GCG | GCC | CTG | AAG | GAC | ATT | GTT | GAT | GAT | GCC | CTT | GGG | GAC | AGG | ATC | TTC | AGC | TTG | GAA | 1086 | |
| G | T | N | K | N | E | T | S | F | G | L | E | M | S | Q | T | G | F | S | S | | 373 | |
| GGC | ACC | AAC | AAG | AAT | GAG | ACC | TCT | TTT | GGG | CTG | GAG | ATG | TCG | CAG | ACT | GGC | TTT | TCT | TCA | | 1146 | |

FIG. 5B

| | | | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|-----|
| H | V | V | E | D | G | I | L | L | G | A | V | G | A | Y | D | W | N | G | A | 393 | |
| CAT | GTG | GTA | GAG | GAT | GGG | ATC | CTG | CTG | GGA | GCT | GTG | GGA | GCC | TAT | GAC | TGG | AAC | GGG | GCG | 1206 | |
| V | L | K | E | T | S | A | G | K | V | I | P | H | R | E | S | Y | L | K | E | 413 | |
| GTG | CTG | AAG | GAG | ACA | AGT | GCT | GGC | AAG | GTG | ATT | CCT | CAC | CGA | GAG | TCC | TAC | CTT | AAG | GAG | 1266 | |
| F | P | E | E | L | K | N | H | A | A | Y | L | G | Y | T | V | T | S | V | V | 433 | |
| TTT | CCA | GAG | GAG | CTG | AAG | AAC | CAT | GCA | GCA | TAC | CTA | GSG | TAC | ACG | GTG | ACG | TCG | GTT | GTG | 1326 | |
| S | S | R | Q | G | G | R | V | Y | V | A | G | A | P | R | F | N | H | T | G | K | 453 |
| TCC | TCC | AGG | CAG | GGG | CGG | GTG | TAT | GTG | GCT | GGA | GCC | CCC | AGA | TTC | AAC | CAT | ACT | GGC | AAG | 1386 | |
| V | I | L | F | S | M | H | N | N | R | S | L | T | I | H | Q | A | L | R | G | 473 | |
| GTC | ATT | CTA | TTC | AGC | ATG | CAC | AAC | AAC | CGG | AGC | CTC | ACC | ATC | CAC | CAG | GCT | CTT | CGG | GGC | 1446 | |
| E | Q | I | G | S | Y | F | G | S | E | I | T | S | V | D | V | N | D | D | R | 493 | |
| GAG | CAG | ATA | GGC | TCC | TAC | TTT | GGG | AGT | GAG | ATT | ACC | TCA | GTG | GAC | GTC | AAT | GAT | GAC | AGA | 1506 | |
| V | T | D | V | L | L | V | G | A | P | M | Y | F | S | E | G | R | E | R | G | 513 | |
| GTG | ACA | GAT | GTG | CTG | CTG | GTG | GGT | GCA | CCC | ATG | TAC | TTC | AGC | GAG | GGC | CGA | GAG | CGA | GGC | 1566 | |
| K | V | Y | V | Y | N | L | R | Q | N | R | F | V | Y | N | G | T | L | K | D | 533 | |
| AAG | GTG | TAT | GTC | TAC | AAC | CTG | AGA | CAG | AAC | CGG | TTT | GTT | TAT | AAT | GGC | ACT | CTG | AAG | GAT | 1626 | |
| S | H | S | Y | Q | N | A | R | F | G | S | C | I | A | S | V | Q | D | L | N | 553 | |
| TCC | CAC | AGC | TAC | CAG | AAC | GCC | CGG | TTC | GGG | TCC | TGC | ATC | GCC | TCT | GTT | CAA | GAC | CTC | AAC | 1686 | |
| Q | D | S | Y | N | D | V | V | V | G | A | P | L | E | D | S | H | R | G | A | 573 | |
| CAA | GAT | TCC | TAC | AAT | GAT | GTG | GTG | GTG | GGG | GCC | CCT | CTG | GAG | GAC | AGC | CAC | AGA | GGG | GCC | 1746 | |
| I | Y | I | F | H | G | F | Q | T | N | I | L | K | K | P | M | Q | R | I | T | 593 | |
| ATC | TAC | ATC | TTC | CAT | GGC | TTC | CAA | ACC | AAT | ATC | CTG | AAG | AAG | CCC | ATG | CAG | AGA | ATA | ACA | 1806 | |
| A | S | E | L | A | P | G | L | Q | H | F | G | C | S | I | H | G | Q | L | D | 613 | |
| GCC | TCG | GAG | CTG | GCT | CCT | GGC | CTG | CAG | CAT | TTT | GGC | TGC | AGC | ATC | CAC | GGA | CAA | CTG | GAC | 1866 | |

FIG. 5C

| | |
|---|------|
| L N E D G L V D L A V G G A L G N A V L | 633 |
| CTC AAT GAG GAT GGG CTT GTG GCA CTA GAC GTC GGC GGC CTG GGC GGT TTG | 1926 |
| W A R P V V Q I N A S L H F E P S K I N | 653 |
| TGG GCC CGC CCT GTA GTT CAA ATC AAC GCC AGC CTC GAG CCT TCC AAG ATC AAC | 1986 |
| I F H K D C K R N G R D A T C L A A F L | 673 |
| ATC TTC CAC AAG GAC TGC AAG CGC AAT GGC AGG GAT GCC ACC TGC CTG GCC TTC CTC | 2046 |
| C F I P I F L A P H F Q T A T V G I R Y | 693 |
| TGC TTC ATA CCC ATC TTC CTG GCA CCC CAC TTC CAA ACA GCA ACC GTC GGC ATC AGG TAC | 2106 |
| N A T M D E R R Y M P R A H L D E G G D | 713 |
| AAT GCA ACC ATG GAT GAG AGG CGG TAT ATG CCG CGG GCA CAT CTG GAC GAG GGC GGA GAC | 2166 |
| Q F T N R A V L L S S G Q E H C Q R I N | 733 |
| CAG TTC ACC AAC AGA GCC GTC CTG CTC TCC TCT GGC CAG GAA CAC TGT CAA AGG ATC AAC | 2226 |
| F H V L D T A D Y V K P V A F S V E Y S | 753 |
| TTC CAT GTC CTG GAC ACT GCA GAC TAC GTG AAG CCA GTG GCC TTC TCC GTG GAG TAC TCC | 2286 |
| L E D P D N G G P M L D N G G W P T T L R V | 773 |
| CTA GAG GAC CCT GAC AAC GGC GGC GAC AAC GGC TGG CCC ACT ACG CTC AGA GTG | 2346 |

FIG. 5D

| | | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| S | V | P | F | W | N | G | C | N | E | D | E | H | C | V | P | D | L | V | L | 793 |
| TCG | GTG | CCC | TTC | TGG | AAT | GGC | TGT | AAT | GAG | GAT | GAA | CAC | TGT | GTC | CCT | GAC | CTT | GTA | CTG | 2406 |
| D | A | R | S | D | L | P | T | A | M | E | Y | C | Q | R | V | L | G | R | P | 813 |
| GAT | GCT | CGG | AGT | GAC | CTG | CCC | ACT | GCC | ATG | GAG | TAC | TGC | CAG | CGG | GTG | CTG | GGG | AGG | CCG | 2466 |
| A | Q | D | C | S | S | Y | T | L | S | F | D | T | T | V | F | I | I | E | S | 833 |
| GCG | CAG | GAC | TGC | TCC | AGC | TAC | ACC | CTG | TCC | TTC | GAC | ACC | ACT | GTC | TTC | ATC | ATA | GAG | AGC | 2526 |
| T | R | R | R | V | A | V | E | A | T | L | E | N | R | G | E | N | A | Y | S | 853 |
| ACG | CGC | CGC | CGG | GTG | GCA | GTG | GAG | GCC | ACG | CTG | GAG | AAC | AGA | GGA | GAG | AAT | GCT | TAC | AGC | 2586 |
| A | V | L | N | I | S | Q | S | E | N | L | Q | F | A | S | L | I | Q | K | D | 873 |
| GCC | GTG | CTC | AAC | ATC | TCC | CAG | TCA | GAG | AAC | CTT | CAG | TTT | GCC | AGC | CTG | ATC | CAG | AAG | GAT | 2646 |
| D | S | D | N | S | I | E | C | V | N | E | E | R | R | L | H | K | K | V | C | 893 |
| GAC | TCA | GAC | AAC | AGC | ATC | GAG | TGT | GTG | AAC | GAG | GAG | AGG | CGG | CTT | CAC | AAG | AAA | GTC | TGC | 2706 |
| N | V | S | Y | P | F | F | R | A | K | A | K | V | A | F | R | L | D | F | E | 913 |
| AAC | GTG | AGC | TAC | CCC | TTC | TTC | AGA | GCC | AAG | GCC | AAG | GTG | GCT | TTC | CGT | CTG | GAT | TTC | GAG | 2766 |
| F | S | K | S | V | F | L | H | H | L | Q | I | H | L | G | A | G | S | D | S | 933 |
| TTC | AGC | AAG | TCT | GTG | TTT | CTG | CAC | CAT | CTT | CAG | ATC | CAT | CTG | GGT | GCC | GGC | AGT | GAC | AGT | 2826 |
| H | E | Q | D | S | T | A | D | D | N | T | A | L | L | R | F | H | L | K | Y | 953 |
| CAC | GAG | CAA | GAC | AGC | ACC | GCA | GAC | GAT | AAC | ACG | GCC | CTC | CTG | CGC | TTC | CAC | CTC | AAA | TAT | 2886 |
| E | A | D | V | L | F | T | R | S | S | S | L | S | H | F | E | V | K | A | N | 973 |
| GAA | GCA | GAC | GTC | CTC | TTT | ACC | AGA | AGC | AGC | AGC | CTG | AGC | CAC | TTC | GAG | GTG | AAG | GCA | AAC | 2946 |
| S | S | L | E | S | Y | D | G | I | G | P | P | F | N | C | V | F | K | V | Q | 993 |
| AGC | TCA | CTT | GAG | AGC | TAT | GAT | GGC | ATC | GGA | CCT | CCG | TTT | AAC | TGT | GTT | TTC | AAG | GTG | CAG | 3006 |
| N | L | G | F | F | P | I | H | G | V | M | M | K | I | T | V | P | I | A | T | 1013 |
| AAC | CTG | GGC | TTT | TTC | CCC | ATC | CAC | GGG | GTG | ATG | ATG | AAG | ATC | ACT | GTG | CCC | ATC | GCC | ACC | 3066 |

FIG. 5E

| | | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| R | G | G | N | R | L | L | M | L | R | D | F | F | T | D | Q | G | N | T | S | 1033 |
| AGG | GGT | GGC | AAC | CGC | CTG | CTG | ATG | TTG | AGG | GAC | TTC | TTC | ACC | GAC | CAG | GGC | AAC | ACA | TCC | 3126 |
| C | N | I | W | G | N | S | T | E | Y | R | S | T | P | T | E | E | D | L | S | 1053 |
| TGT | AAC | ATC | TGG | GGG | AAC | AGC | ACA | GAG | TAC | AGG | AGT | ACC | CCA | ACT | GAG | GAA | GAT | CTG | AGC | 3186 |
| H | A | P | Q | R | N | H | S | N | S | D | V | V | S | I | I | C | N | L | R | 1073 |
| CAT | GCC | CCA | CAG | AGG | AAT | CAC | AGC | AAC | TCC | GAC | GTG | GTC | TCC | ATC | ATC | TGC | AAC | CTG | AGG | 3246 |
| L | A | P | S | Q | E | T | S | F | Y | L | V | G | N | L | W | L | T | S | L | 1093 |
| CTG | GCC | CCC | AGC | CAG | GAG | ACC | AGC | TTT | TAC | CTG | GTG | GGA | AAC | CTG | TGG | CTG | ACA | TCT | CTG | 3306 |
| K | A | L | K | Y | R | S | L | K | I | T | V | N | A | A | L | Q | R | Q | F | 1113 |
| AAA | GCA | CTC | AAG | TAC | AGG | TCT | CTG | AAG | ATC | ACA | GTC | AAC | GCC | GCC | TTG | CAG | AGG | CAG | TTC | 3366 |
| H | S | P | F | I | F | R | E | E | D | P | S | R | Q | V | T | F | E | I | S | 1133 |
| CAC | AGC | CCC | TTC | ATC | TTC | CGA | GAG | GAA | GAC | CCT | AGC | CGC | CAG | GTC | ACA | TTT | GAG | ATC | TCC | 3426 |
| K | Q | E | D | W | Q | V | P | I | W | I | I | V | G | S | T | L | G | G | L | 1153 |
| AAG | CAA | GAA | GAC | TGG | CAG | GTC | CCC | ATC | TGG | ATC | ATC | GTG | GGC | AGC | ACT | CTG | GGG | GGC | CTC | 3486 |
| L | L | L | A | L | L | V | L | A | L | W | K | L | G | F | F | K | S | A | K | 1173 |
| TTG | CTG | CTG | GCC | CTG | CTG | GTC | CTG | GCA | CTG | TGG | AAG | CTC | GGT | TTT | TTT | AAA | AGT | GCC | AAG | 3546 |
| R | K | R | E | P | G | L | G | P | I | P | K | E | L | K | * | | | | | 1189 |
| CGC | AAG | AGG | GAA | CCT | GGC | CTA | GGC | CCC | ATC | CCC | AAA | GAG | CTC | AAG | TGA | | | | | 3594 |

FIG. 5F

GGATCCAGAGGAAGCTTCAAGTTGTTGGGGGCCCCGACACCAAGTCCGGGGGTGTGCAAGGCCCGGGCTATGGCCTCAGAGC 3673
TGGAGCAAGAGGACACCCATGGACTTTACATTGAGTTTCATCTCTGGAGCAATGGCAACTGATCGCTGTGGGATGGACC 3752
TCCAGCCCGTGTTCGGAGTGACTTCTCTACTAGAGGCCAAGACCTCCAGTACAGATCCCCTAAGGACTAAAGGGACCCCT 3831
CCAAGACCCACCGGTGGCCTTCCCCCAAGGCTCTGTGGAGGGAATCTGCTGCCCCCAACACTAAGGTGCTAGGGTTTGTG 3910
TTGTCTTGTCCCACTCTGAAGAACTCAAGGGGAAGCTAGCAAAATACGAGCCCCACTGTGCACGCCCTATCTGGACCTAT 3989
AGCTTCAGGACTCCATGGCCCCGGGAGCATAACTCCGGCCGACCCCTTACCCCTTGCTCTTGGTTCCCCAACACAGCCA 4068
TCCCTTCTCCACAGATAGACCCCAAGATCTCTATGAGTGACCCAGAGTTGTGTGGCTTGATATGCCAGCTGCAGGCCA 4147
GTGATGGAAACACATCCAGGGACTGTGCCAAATGAACTGGAGGGAATCAGGATGCTGGGTACCTTCCAAACATCAGAAA 4226
GCAGCAGGGCTCTCCAGCAGTAACCAACCCACCTCAATCTGCTCTCCAGCCTCCCCCGCAGCTCCTCTCAGCCGCTT 4305
TGGTTGCTAGGCAACTCGTGCTAGGGAGCCTGAGGAGGATGATAAAAATCCGTGGCTCTCCGTACCCCTCTTGTCTGACA 4384
ACTGCAGCTTGTCCCTGGTAGATGCACTGTGATAGGGAGCCTCTGTTTTCTTACATCAGCACCTACACACACACACAC 4463
ACACACACACACACACACACACACACACGCAACGCACTCCCAACCATCAGCGTGTGGGAGAAGGGCTCCTGTGA 4542
GGGGGGCGCTGTGGGGGGCCTTCTAGGCATGCACTGAACAAAGCCCCATGCAAGGCCCTTGGGAACCTTAAACCTCAGAGG 4621
CAGATAGGCATCTCGCCCGCCAGCTCTGGGATAAGCCCCAAGGACAAAGGGCATGCCGGATGCCCAGCAGAGTAATTTA 4700
TGCCCTTAACCTTTTGTGTGAGATGGAATGGACACATCAGTGTCCCCCTGTTCATCTGTTCATCTGAAACAAATGTCTGT 4779
TCCCCCTATGCGGTACGACCTTTACTGTGAGAAATATATTAAAGAAAATGTTCTTAAAAAATAAAAAAAGGGGCGCCGC 4858

FIG. 5G

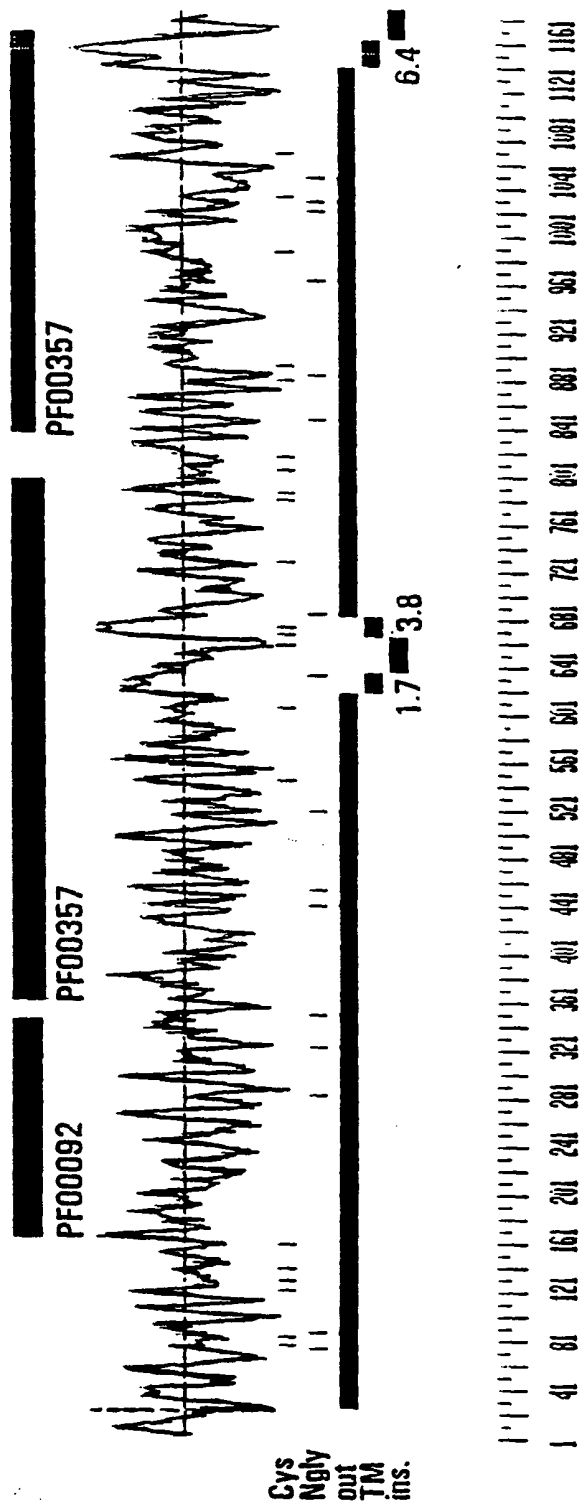


FIG. 6

```

      10      20      30      40      50      60      7
inputs  GTCGACCCACGCGTCCGGGCGCCCCGCGCAGTGCAGGCTGCAGGCGCCGCGCCGAGGAGGCTGCCGCTC
      :
      G-----

      80      90      100      110      120      130      14
inputs  GGCTTGCCGCCCCCGCCGCTGCACACCCGACCCAGCCGCGCTGCCGCGGGCCATGGACCTGCCC/
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      -----CGACACCCGCGCCGC-----GCAGCGGGCCATGGACTTTCCC/
      10      20      30      40

      150      160      170      180      190      200      21
inputs  GGGCCTGGTGGTGGCCTGGGCGCTCAGCCTGTGGCCAGGGTTCACGGACACCTTCAACATGGACACCAC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      GGGCCTCCTGGTGGCCTGGACCTCAGCCTTTGGCCAGGATTCACGGACACCTTCAATATGGATACCA/
      50      60      70      80      90      100      110

      220      230      240      250      260      270      28
inputs  AAGCCCCGGGTATCCCTGGCTCCAGGACCGCCTTCTTTGGCTACACAGTGCAGCAGCAGCAGCATCAG/
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      AATCCCCGGGTATTGCTGGTCCCAGTGCCTTCTTTGGCTACACAGTACAGCAGCATGATATCAG/
      120      130      140      150      160      170      180

      290      300      310      320      330      340      35
inputs  GCAATAAGTGGCTGGTCTGTTGGGCGCCCCACTGGAAACCAATGGCTACCAGAAGACGGGAGACGTGTAC/
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      GCAAGAAGTGGTTGGTTGTGGGTGCCCAATGGAAACCAATGGCCATCAGAAGACAGGAGACGTATAC/
      190      200      210      220      230      240      250

      360      370      380      390      400      410      42
inputs  GTGTCCAGTGATCCACGGGAAGTGCACCAAAGTCAACCTGGGAAGGGTCAACCTGTCCAACGTGTCCG/
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      ATGTCCGGTAACCCAGGGCAAGTGCACCAAGCTCAACCTGGGCAGGGTCACTCTGTCCAATGTGTCTG/
      260      270      280      290      300      310      320

      430      440      450      460      470      480      49
inputs  CGGAAAGACAACATGCGCCTCGGCCTTAGTCTCGCCACCAACCCCAAGGACAACAGCTTCCTGGCCTG/
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      AGGAAGGACAACATGCGCCTCGGCCTGAGCCTTGCCACCAACCCCAAGGACAACAGCTTCCTGGCCTG/
      330      340      350      360      370      380      390

      500      510      520      530      540      550      56
inputs  GCCCCCTCTGGTCTCATGAGTGTGGGAGCTCCTACTACACCACAGGGATGTGTTCAAGAGTCAACTCC/
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      GCCCTCTCTGGTGCACAGAGTGTGGAAGCTCCTACTACACCACTGGCATGTGCTCACGGGTCAACTCC/
      400      410      420      430      440      450      460

      570      580      590      600      610      620      63
inputs  CTTCAGGTTCTCCAAGACCGTGGCCCCAGCTCTCCAAAGGTGCCAGACCTACATGGACATCGTCATTG/
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      CTTCAGATTCTCTAAGACGGTAGCTCCGGCACTTCAGAGGTGCCAGACTTACATGGACATCGTCATTG/
      470      480      490      500      510      520      530

      640      650      660      670      680      690      70
inputs  CTGGATGGCTCCAACAGCATCTACCCCTGGGTGGAGGTTCAAGCACTTCCTCATCAACATCCTGAAAAA/
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      TTAGATGGCTCCAACAGCATCTACCCCTGGGTGGAGGTTCAACACTTCCTCATCAATATCCTCAAAAAA/
      540      550      560      570      580      590      600

```

FIG. 7A


```

      1480      1490      1500      1510      1520      1530      1540
inputs AACCACACGGGCAAGGTCATCCTGTTTACCATGCACAACAACCGGAGCCTCACCATCCACCAGGCTATGC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      AACCATACTGGCAAGGTCATTCTATTTCAGCATGCACAACAACCGGAGCCTCACCATCCACCAGGCTCTTC
      1380      1390      1400      1410      1420      1430      1440

      1550      1560      1570      1580      1590      1600      1610
inputs GGGGCCAGCAGATAGGCTCTTACTTTGGGAGTGAAATCACCTCGGTGGACATCGACGGCGACGGCGTGAC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      GGGGCCAGCAGATAGGCTCCTACTTTGGGAGTGAGATTACCTCAGTGGACGTCAATGATGACAGAGTGAC
      1450      1460      1470      1480      1490      1500      1510

      1620      1630      1640      1650      1660      1670      1680
inputs TGATGTCTGCTGGTGGGCGCACCCATGTACTTCAACGAGGGCCGTGAGCGAGGCAAGGTGTACGTCTAT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      AGATGTGCTGCTGGTGGGTGCACCCATGTACTTCAGCGAGGGCCGAGCGAGGCAAGGTGTATGTCTAC
      1520      1530      1540      1550      1560      1570      1580

      1690      1700      1710      1720      1730      1740      1750
inputs GAGCTGAGACAGAACCGGTTTGTATTATAACGGAACGCTAAAGGATTACACAGTTACCAGAAATGCCCGAT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      AACCTGAGACAGAACCGGTTTGTATTATAATGGCACTCTGAAGGATTCCACAGCTACCAGAACGCCCGGT
      1590      1600      1610      1620      1630      1640      1650

      1760      1770      1780      1790      1800      1810      1820
inputs TTGGGTCCTCCATTGCCTCAGTTCGAGACCTCAACCAGGATTCTTACAATGACGTGGTGGTGGGAGCCCC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      TCGGGTCCTGCATCGCCTCTGTTCAAGACCTCAACCAAGATTCTTACAATGATGTGGTGGTGGGAGCCCC
      1660      1670      1680      1690      1700      1710      1720

      1830      1840      1850      1860      1870      1880      1890
inputs CCTGGAGGACAACCACGCGAGGACCATCTACATCTTCCACGGCTTCCGAGGCAGCATCCTGAAGACACCT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      TCTGGAGGACAGCCACAGAGGGGCCATCTACATCTTCCATGGCTTCCAAACCAATATCCTGAAGAAGCCC
      1730      1740      1750      1760      1770      1780      1790

      1900      1910      1920      1930      1940      1950      1960
inputs AAGCAGAGAATCACAGCCTCAGAGCTGGCTACCGGCCTCCAGTATTTTGGCTGCAGCATCCACGGGCAAT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      ATGCAGAGAATAACAGCCTCGGAGCTGGCTCCTGGCCTGCAGCATTTTGGCTGCAGCATCCACGGACAAC
      1800      1810      1820      1830      1840      1850      1860

      1970      1980      1990      2000      2010      2020      2030
inputs TGGACCTCAATGAGGATGGGCTCATCGACCTGGCAGTGGGAGCCCTTGGCAACGCTGTGATTCTGTGGTC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      TGGACCTCAATGAGGATGGGCTTGTGGACCTAGCAGTGGGCGCCCTGGGCAACGCTGTGGTTTGTGGGC
      1870      1880      1890      1900      1910      1920      1930

      2040      2050      2060      2070      2080      2090      2100
inputs CCGCCCAGTGGTTTCAGATCAATGCCAGCCTCCACTTTGAGCCATCCAAGATCAACATCTTCCACAGAGAC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      CCGCCCTGTAGTTCAAATCAACGCCAGCCTCCACTTTGAGCCTTCCAAGATCAACATCTTCCACAAGGAC
      1940      1950      1960      1970      1980      1990      2000

      2110      2120      2130      2140      2150      2160      2170
inputs TGCAAGCGCAGTGGCAGGGATGCCACCTGCCTGGCCGCTTCTCTGCTTCACGCCCATCTTCTGGCAC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      TGCAAGCGCAATGGCAGGGATGCCACCTGCCTGGCCGCTTCTCTGCTTCATACCCATCTTCTGGCAC
      2010      2020      2030      2040      2050      2060      2070

      2180      2190      2200      2210      2220      2230      2240
inputs CCCATTTCCAAACAACAACCTGTTGGCATCAGATACAACGCCACCATGGATGAGAGGCGGTATACACCGAC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      CCCACTTCCAAACAGCAACCGTCTGGCATCAGGTACAATGCAACCATGGATGAGAGGCGGTATATGCCGCC
      2080      2090      2100      2110      2120      2130      2140

```

FIG. 7C


```

      2950      2960      2970      2980      2990      3000      3010
inputs ACCAAGGAAGACAACGTGGCCCCCTTACGCTTCCACCTCAAATACGAGGCTGACGTCTTCCACCAGGA
      2850      2860      2870      2880      2890      2900      2910
      3020      3030      3040      3050      3060      3070      3080
inputs GCAGCAGCCTGAGCCACTACGAGGTCAAGCCCCAACAGCTCGCTGGAGAGATACGATGGTATCGGGCCTCC
      2920      2930      2940      2950      2960      2970      2980
      3090      3100      3110      3120      3130      3140      3150
inputs CTTCAGCTGCATCTTCAGGATCCAGAACTTGGGCTTGTTCCTCCATCCACGGGATGATGATGAAGATCACC
      2990      3000      3010      3020      3030      3040      3050
      3160      3170      3180      3190      3200      3210      3220
inputs ATTCCCATCGCCACCAGGAGCGGCAACCGCCTACTGAAGCTGAGGGACTTCCTCACGGACGAGGCGAACA
      3060      3070      3080      3090      3100      3110      3120
      3230      3240      3250      3260      3270      3280      3290
inputs CGTCCTGTAACATCTGGGGCAATAGCACTGAGTACCGGCCACCCAGTGGAGGAAGACTTGCGTCGTGC
      3130      3140      3150      3160      3170      3180      3190
      3300      3310      3320      3330      3340      3350      3360
inputs TCCACAGCTGAATCACAGCAACTCTGATGTCGTCTCCATCAACTGCAATATACGGCTGGTCCCCAACCCAG
      3200      3210      3220      3230      3240      3250      3260
      3370      3380      3390      3400      3410      3420      3430
inputs GAAATCAATTTCCATCTACTGGGGAACCTGTGGTTGAGGTCCCTAAAAGCACTCAAGTACAAATCCATGA
      3270      3280      3290      3300      3310      3320      3330
      3440      3450      3460      3470      3480      3490      3500
inputs AAATCATGGTCAACGCAGCCTTGCAGAGGCAGTTCCACAGCCCCTTCATCTTCCGTGAGGAGGATCCAC
      3340      3350      3360      3370      3380      3390      3400
      3510      3520      3530      3540      3550      3560      3570
inputs CCGCCAGATCGTGTGAGATCTCCAAGCAAGAGGACTGGCAGGTCCCATCTGGATCATTTGATGGCAGC
      3410      3420      3430      3440      3450      3460      3470
      3580      3590      3600      3610      3620      3630      3640
inputs ACCCTGGGGGGCCTCTACTGCTGGCCCTGCTGGTCTGGCACTGTGGAAGCTCGGCTTCTTTAGAAGTG
      3480      3490      3500      3510      3520      3530      3540
      3650      3660      3670      3680      3690      3700      3710
inputs CCAGGCGCAGGAGGGAGCCTGGTCTGGACCCACCCCAAGTCTGGAGTGAGGCTCCAGAGGAGACTT
      3550      3560      3570      3580      3590      3600      3610

```

FIG. 7E

```

      3720      3730      3740      3750      3760      3770      378
inputs TGAGTTGATGGGGGCCAGGACACCAGTCCAGGTAGTGTGAGACCCAGGCCTGTGGCCCCACCGAGCTC
      3620      3630      3640      3650      3660      3670
      3790      3800      3810      3820      3830      3840      385
inputs AGCGGAGAGGAAGCCAGCTGGCTTTGCACTTGACCTCATCTCCCGAGCAATGGCGCCTGCTCCCTCCAG
      3680      3690      3700      3710      3720      3730      3740
      3860      3870      3880      3890      3900      3910      392
inputs ATGGAAGCTCAAGCTGGTTTAAAGTGAAGTGGCCCTACTGGGAGACTGGGACACCTTTAACACAGACCCC
      3750      3760      3770      3780      3790      3800      3810
      3930      3940      3950      3960      3970      3980
inputs AGGGATTAAAGGGACACCCCTACACACACCCAGGCCACGCCAAGGCCT-CCCTCA-GGCTCTGTGG?
      3820      3830      3840      3850      3860      3870
      3990      4000      4010      4020      4030      4040      4050
inputs GGCATTGTCTGCCCCAGCTACTAAGGTGCTAGGAATTCGTAATCATCCCCATCTCCAGAGAAACCCAC
      3880      3890      3900      3910      3920      3930      394
      4060      4070      4080      4090      4100      4110      4120
inputs GAGGAAGACTGTAAATACGAACCCAATCTGCACACTCCAGGCCTCTAGTTCCAGAAGGATCCAAGACA?
      3950      3960      3970      3980      3990      4000
      4130      4140      4150      4160      4170      4180      4190
inputs ACAGATCTGAATCTGCCCTTTTCTCTACCCATCCACCCCTCCA-TTGGCTCCCAAGTCACACCCAC
      4010      4020      4030      4040      4050      4060
      4200      4210      4220      4230      4240      4250      4260
inputs CCCTTCCCCATAGATAGGCCCCTGGGGCTCCTGAAGAATGAACCCAAGAGCAAGGGCTTGATGGTGAC?
      4070      4080      4090      4100      4110      4120      4130
      4270      4280      4290      4300      4310      4320      4330
inputs CTGCAAGCCAGGGATGAAGAAAGACTCTGAGATGTGGAGACTGATGGCCAGGCAAGTGGGACCAGGATA?
      4140      4150      4160      4170      4180      4190      4200
      4340      4350      4360      4370      4380      4390      4400
inputs TGGACGCTGTCTTGAGATGAGAG-GTAGCCGGGCTCTGCACCCACGTGCATTCACATTGACCCGCAACTC
      4210      4220      4230      4240      4250

```

FIG. 7F

```

4410      4420      4430      4440      4450      4460      4470
inputs CACATTCC--CCCACCAGCTGCAGCCCCCTTGCTCTCAGCTGCCAACCCTCCCGGGTCACCTTTTGTTC
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CTCAATCCTGCTCTCCAGC--CT-CCCC-----CGCAGCT-----CCTCTCAG-CCGCTTTGGTTG
4260      4270      4280      4290      4300      4310

4480      4490      4500      4510      4520      4530
inputs GGTACCTCATG---GGAAGC---ATGTGGATGACACAATCCCTGGGGCTGTGCATTCCCAC--GTCT
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
GGCAACTCGTGTCTAGGGAGCCTGAGGAGGATGATAAAAATCC-GTGGCTCTCCGTACCCCTCTTGCTC
4320      4330      4340      4350      4360      4370      4380

4540      4550      4560      4570      4580      4590      4600
inputs TGCTGCAGCCTGCCCTAGACATGGACGCACTGGCCTGGCTGCAGCTGGGCAGCAGGGGTAGGGGTAC
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
AACTGCAGCTTGTCCTGG---TAGATGCACTG---TGATAG---GGAGCCTCTGTTCTTACATCA
4390      4400      4410      4420      4430      4440

4610      4620      4630      4640      4650      4660      4670
inputs AGCCTCCCCTCCCTGTATCACCCCCTCCCTACACACACACACACACACACACACACT-GCCTCC
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
A-CCTACACACACACACACACACACAC-ACACACACACACACACACACACACACGACGACCTCC
4450      4460      4470      4480      4490      4500      451

4680      4690      4700      4710      4720      4730      4740
inputs TCCTTCCCTCATGCCCGC-CAGTGCACAGGGAAGGGCTTGCCAGCGCTGTTGAGGGGTCCCCTCTGG
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ACCATCA-GCGTGTGGGAGAAGGGCTCCTGTAGGGG-----GCGCTGTGGGGG---CCTTCTAG
4520      4530      4540      4550      4560

4750      4760      4770      4780      4790      4800      4810
inputs ATGCACTGAATAAAGCACGTGCAAGGACTCCCGGAGCCTGTGCAGCCTTGGTGGCAAATATCTCATCT
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ATGCACTGAACAAAGCCCATGCAAGGCCCTGGGAACCTTA--AACCTCAGAGGCAGATAG-GCATCT
4570      4580      4590      4600      4610      4620      4630

4820      4830      4840      4850      4860      4870      4880
inputs CCGGCCCCCAGGACAAGTGGTATGACCAGTGATAATGCCCCAAGGACAAGGGGCGTGCTGGCGCCCA
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CCGGCC-----AGCT-----CTGGGATAA-GCCCCAAGGACAAGGGGCATGCCGGATGCCCA
4640      4650      4660      4670      4680

4890      4900      4910      4920      4930      4940
inputs GGAGTAATTTATGCCTTAGTCTTGTGTTTGGAGGTAGAAATG----CAAGGGGGACACATGAAAGGCATC
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
AGAGTAATTTATGCCTTAACCTTTTGTGAGATGGAAATGGACACATCAGTGTCCCCTGT---CATCT
4690      4700      4710      4720      4730      4740      4750

4950      4960      4970      4980      4990      5000      501
inputs TCCCC-CTGTGCATAGT--ACGA----CCTTTACTGT-CGTATTTTTGAAAAATTAAAAATACAGTGTT
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
TTTCATCTGAACAATGTCATCTGTTCCCCTATGCGGTACGACCTTTACTGTTCAGAAATATATTTAAGAA
4760      4770      4780      4790      4800      4810      4820

5020      5030      5040
inputs AAAAAC-AAAAAAAAAAAAAAAAAGGGCGGCCGC
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ATGTTCTAAAAAAAAAAAAAAAAAGGGCGGCCGC
4830      4840      4850

```

FIG. 7G

| | | | | | | | |
|--------|----------------|--------------|---------------|--------------|-----------|---------------|-----------------|
| | 10 | 20 | 30 | 40 | 50 | 60 | 70 |
| inputs | MDLPRGLVVAWALS | SLWPGFTDTFNM | DTKPRVIPGSRTA | FFGYTVQQHDIS | GNKWL | VVGAPLETNGYQK | |
| | MDLPRGLVVAWALS | SLWPGFTDTFNM | DTKPRVIPGSRTA | FFGYTVQQHDIS | GNKWL | VVGAPLETNGYQK | |
| | 10 | 20 | 30 | 40 | 50 | 60 | 70 |
| | 80 | 90 | 100 | 110 | 120 | 130 | 140 |
| inputs | GDVYKCPVIHGNCT | KLNLGRVTLSNV | SERKDNMRLGL | SLATNPKDNS | FLACSPLWS | HECGSSYYTTGMC | |
| | GDVYKCPVTQGNCT | KLNLGRVTLSNV | SERKDNMRLGL | SLATNPKDNS | FLACSPLWS | HECGSSYYTTGMC | |
| | 80 | 90 | 100 | 110 | 120 | 130 | 140 |
| | 150 | 160 | 170 | 180 | 190 | 200 | 210 |
| inputs | RVNSNFRFSKTVAP | ALQRCQTYMDI | VIVLDGSNSI | YPWVEVQHFL | INILKKFYI | GPGQIQVG | VVQYGED |
| | RVNSNFRFSKTVAP | ALQRCQTYMDI | VIVLDGSNSI | YPWVEVQHFL | INILKKFYI | GPGQIQVG | VVQYGED |
| | 150 | 160 | 170 | 180 | 190 | 200 | 210 |
| | 220 | 230 | 240 | 250 | 260 | 270 | 280 |
| inputs | VHEFHLNDYRSVK | DVVEAASHIEQ | RGGTETR | TAFGIEFAR | SEAFQKGR | KGAKKVMIV | ITDGESHDSPD |
| | VHEFHLNDYRSVK | DVVEAASHIEQ | RGGTETR | TAFGIEFAR | SEAFQKGR | KGAKKVMIV | ITDGESHDSPD |
| | 220 | 230 | 240 | 250 | 260 | 270 | 280 |
| | 290 | 300 | 310 | 320 | 330 | 340 | 350 |
| inputs | EKVIQQSERDNV | TRYAVAVLGY | NNRRGINPET | FLNEIKYI | ASDPDDKH | FFNVTDEA | ALKDIVDALGDRI |
| | EKVIQQSERDNV | TRYAVAVLGY | NNRRGINPET | FLNEIKYI | ASDPDDKH | FFNVTDEA | ALKDIVDALGDRI |
| | 290 | 300 | 310 | 320 | 330 | 340 | 350 |
| | 360 | 370 | 380 | 390 | 400 | 410 | 420 |
| inputs | SLEGTNKNETS | SFGLEMSQTG | FSHVVEDG | VLLGAVGAY | DWNGAVL | KETSAGKVI | PLRESYLKEFPEELK |
| | SLEGTNKNETS | SFGLEMSQTG | FSHVVEDG | VLLGAVGAY | DWNGAVL | KETSAGKVI | PLRESYLKEFPEELK |
| | 360 | 370 | 380 | 390 | 400 | 410 | 420 |
| | 430 | 440 | 450 | 460 | 470 | 480 | 490 |
| inputs | HGAYLGYTVTS | VVSSRQGRVY | VAGAPRNFHT | GKVLFTMH | NNRSLTIHQ | AMRGOQIG | SYFGSEITSVDI |
| | HGAYLGYTVTS | VVSSRQGRVY | VAGAPRNFHT | GKVLFTMH | NNRSLTIHQ | AMRGOQIG | SYFGSEITSVDI |
| | 430 | 440 | 450 | 460 | 470 | 480 | 490 |
| | 500 | 510 | 520 | 530 | 540 | 550 | 560 |
| inputs | GDGVTDVLLV | GAPMYFNEGR | ERGKVVYEL | RQNRVYNG | TLKDSHSY | QNAFSGSS | IASVRDLNQDSYN |
| | GDGVTDVLLV | GAPMYFNEGR | ERGKVVYEL | RQNRVYNG | TLKDSHSY | QNAFSGSS | IASVRDLNQDSYN |
| | 500 | 510 | 520 | 530 | 540 | 550 | 560 |
| | 570 | 580 | 590 | 600 | 610 | 620 | 630 |
| inputs | VVGAPLEDNH | GAIYIFHG | FRGSILKTP | KQRITASE | LATGLQY | FGCSIHG | QDLNEDGLID |
| | VVGAPLEDNH | GAIYIFHG | FRGSILKTP | KQRITASE | LATGLQY | FGCSIHG | QDLNEDGLID |
| | 570 | 580 | 590 | 600 | 610 | 620 | 630 |
| | 640 | 650 | 660 | 670 | 680 | 690 | 700 |
| inputs | VILWSRPVVQ | INASLHFE | PSKINIFH | RDCRSGR | DATCLAA | FLCFPI | IFLAPHFQT |
| | VILWSRPVVQ | INASLHFE | PSKINIFH | RDCRSGR | DATCLAA | FLCFPI | IFLAPHFQT |
| | 640 | 650 | 660 | 670 | 680 | 690 | 700 |

FIG. 8A

| | | | | | | | |
|--------|---|------|------|------|------|------|-----|
| | 710 | 720 | 730 | 740 | 750 | 760 | 77 |
| inputs | RYTPRAHLDEGGDRFTNRAVLLSSGQELCERINFHVLDTADYVKPVTFVSVEYSLEDPDHGPMDDGWPI | | | | | | |
| | :: :: | | | | | | |
| | RYMPRAHLDEGGDQFTNRAVLLSSGQEHQCRINFHVLDTADYVKPVAFVSVEYSLEDPDNGPMLDNGWPI | | | | | | |
| | 710 | 720 | 730 | 740 | 750 | 760 | 77 |
| | 780 | 790 | 800 | 810 | 820 | 830 | 84 |
| inputs | LRVSVFPFWNGCNEDEHCVLDLVDARS DLPTAMEY CQRVLRKPAQDCSAYTLSFDTTVFIIESTRQRVZ | | | | | | |
| | :: :: | | | | | | |
| | LRVSVFPFWNGCNEDEHCVLDLVDARS DLPTAMEY CQRVLRKPAQDCSSYTLSFDTTVFIIESTRRRVZ | | | | | | |
| | 780 | 790 | 800 | 810 | 820 | 830 | 84 |
| | 850 | 860 | 870 | 880 | 890 | 900 | 91 |
| inputs | EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIIECVNEERRLQKQVCNVSYPPFFRAKAKVAFR | | | | | | |
| | :: :: | | | | | | |
| | EATLENRGENAYSAVLNISQSEN LQFASLIQKDDSDNSIECVNEERRLHKVCNVSYPPFFRAKAKVAFR | | | | | | |
| | 850 | 860 | 870 | 880 | 890 | 900 | 91 |
| | 920 | 930 | 940 | 950 | 960 | 970 | 98 |
| inputs | DFEFSKSIFLHLEIELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLFTRSSSLSHYEVKPNSSLERY | | | | | | |
| | :: :: | | | | | | |
| | DFEFSKSVFLHHLQIHLGAGSDSHEQDSTADDNTALLRFHLKYEADVLFTRSSSLSHFEVKANSSLESY | | | | | | |
| | 920 | 930 | 940 | 950 | 960 | 970 | 98 |
| | 990 | 1000 | 1010 | 1020 | 1030 | 1040 | 105 |
| inputs | GIGPPFSCIFRIQNLGLFPIHGMMMKITIP IATRSGNRLLKLRDFLTDEANTSCNIWGNSTEYRPTPVE | | | | | | |
| | :: :: | | | | | | |
| | GIGPPFNCVFKVQNLGFFPIHGVMK I TVPIATRGGNRLLMLRDFFTDQNTSCNIWGNSTEYRSTPTE | | | | | | |
| | 990 | 1000 | 1010 | 1020 | 1030 | 1040 | 105 |
| | 1060 | 1070 | 1080 | 1090 | 1100 | 1110 | 112 |
| inputs | DLRRAPQLNHSNSDVVSINCNIRLVPNQEI NFHLLGNLWLRSLKALKYKSMKIMVNAALQRFHSPFIF | | | | | | |
| | :: :: | | | | | | |
| | DLSHAPQRNHSNSDVVSIIICNLRLAPSQETS FYLVGNLWLTSLKALKYRSLKITVNAALQRFHSPFIF | | | | | | |
| | 1060 | 1070 | 1080 | 1090 | 1100 | 1110 | 112 |
| | 1130 | 1140 | 1150 | 1160 | 1170 | 1180 | |
| inputs | EEDPSRQIVFEISKQEDWQVPIWIIVGSTLGGLLLLALLVLALWKLGF FR SARRRREPGLDPTPKVLE | | | | | | |
| | :: :: | | | | | | |
| | EEDPSRQVTFEISKQEDWQVPIWIIVGSTLGGLLLLALLVLALWKLGF FKS AKRREPGLGPIPKELK | | | | | | |
| | 1130 | 1140 | 1150 | 1160 | 1170 | 1180 | |

FIG. 8B

```

int_alpha_rpt: domain 1 of 5, from 37 to 90: score 29.5, E = 7.6e-05
      *-pgsyFGysvagvgDlngDksGypDl1lvGAPr.....GaVYvyf
      ++FGy+v   +D+   G  + l+vGAP+++++ +++G VY+++
T259   37   RTAFFGYTVQ-QHDIS----GNKW-LVVGAPLetngyqktGDVYKCP 77

      gssnsgggrciplqnls<-*
      +       g+c+ l   +
T259   78 VI----HGNCTKLNLR      90

```

Figure 9A

```

>int_alpha_rpt: domain 2 of 5, from 421 to 472: score 19.2, E = 0.097
      *-pgsyFGysvagvgDlngDksGypDl1lvGAPr....GaVYvyfgssn
      +g+y+Gy+v++v+      +++++GAPr +++G+V ++  +
T259   421   HGAYLGyTVTSVVS-----RQGRVYVAGAPRfnhtGKVILFTMHN- 461

      sgggrciplqnls<-*
      +++++ q  +
T259   462 --NRSLTIHQAMR      472

```

Figure 9B

```

>int_alpha_rpt: domain 3 of 5, from 476 to 532: score 48.4, E = 1.5e-10
      *-pgsyFGysvagvgDlngDksGypDl1lvGAPr.....GaVYvyf
      +gsyFG+ ++  +D++gD  G +D+llvGAP+ ++++++G+VYvy
T259   476   IGSYFGSEIT-SVDIDGD--GVTDVLLVGAPMyfnegrerGKVYVYE 519

      gssnsgggrciplqnls<-*
      +       +r+ +  +l
T259   520 LR----QNRVYNGTLK      532

```

Figure 9C

```

>int_alpha_rpt: domain 4 of 5, from 538 to 593: score 62.6, E = 8.4e-15
      *-pgsyFGysvagvgDlngDksGypDl1lvGAPr.....GaVYvyfgss
      +++++FG+s+a+v+Dln+D  +y+D ++vGAP+++++ Ga+Y+++g +
T259   538   QNARFGSSIASVRDLNQD--SYND-VVVGAPLednhaGAIYIFHGFR 581

      nsggggrciplqnl<-*
      +  +  +p q ++
T259   582 G--SILKTPKQRIT      593

```

Figure 9D

```

>int_alpha_rpt: domain 5 of 5, from 600 to 654: score 16.5, E = 0.57
      *-pgsyFGysvagvgDlngDksGypDl1lvGAPr.....GaVYvyfgss
      yFG s+ g Dln D G D l+vGA ++      +v+++++
T259   600   GLQYFGCSIHGQLDLNED--GLID-LAVGALGnavilWSRPVVQINA 643

      nsggggrciplqnl<-*
      +++++p+ + +
T259   644 ---SLHFEPKINI      654

```

Figure 9E

Figure 10A

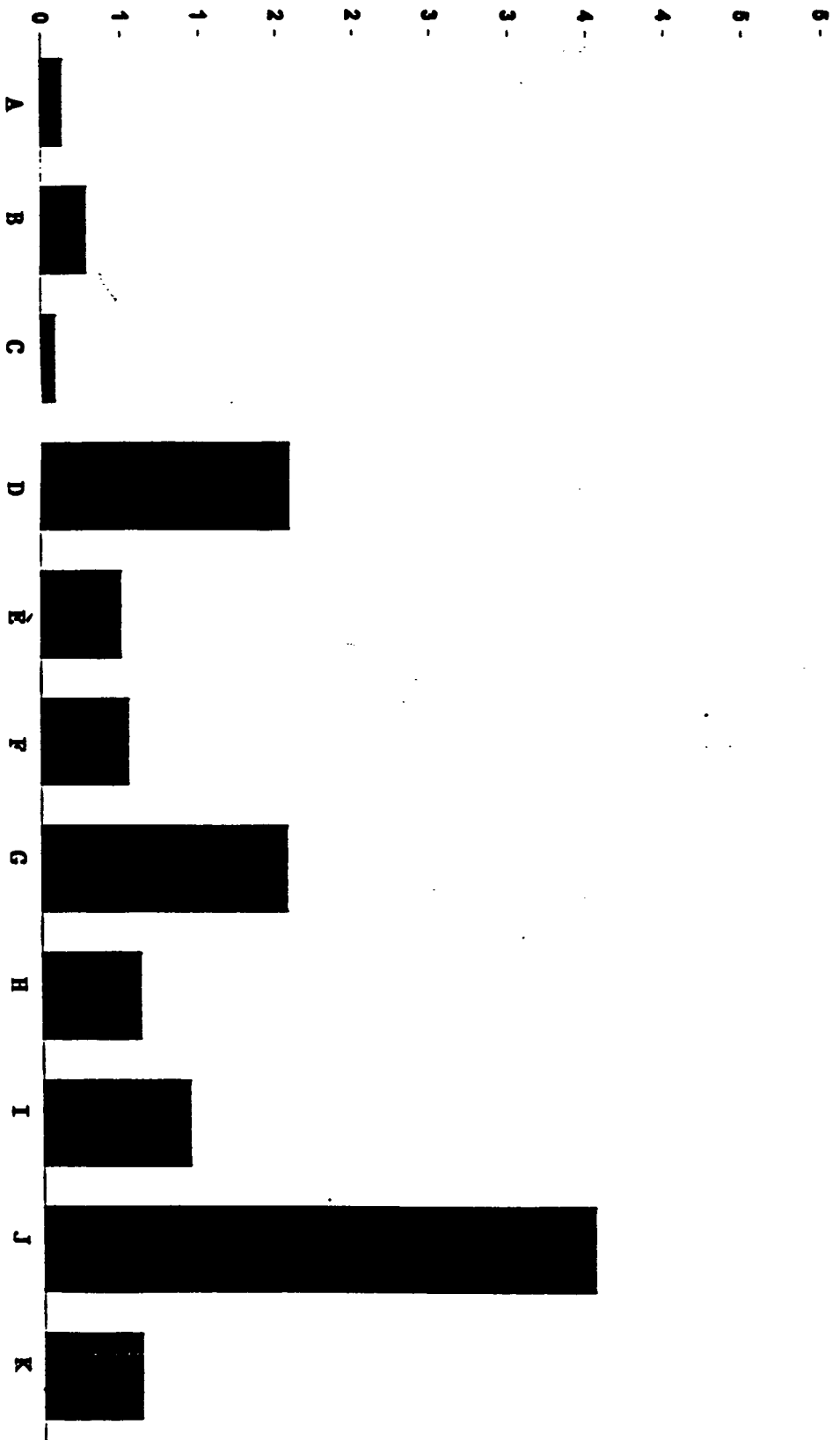


Figure 10B

